보건학석사 학위논문

최소 사망온도의 모수적 추정 : 과산포자료에서 비선형 모수를 포함한 일반화 선형모형을 이용하여

Parametric Estimation of Minimum Mortality Temperature with Over-dispersion: Using Non-linear Generalized Linear Model

2015년 6월

서울대학교 보건대학원

보건학과 보건통계학전공

이환희

최소 사망온도의 모수적 추정
: 과산포자료에서 비선형 모수를 포함한 일반화 선형모형을 이용하여

지도교수 김 호

이 논문을 보건학 석사학위논문으로 제출함

2015년 5월

서울대학교 보건대학원

보건학과 보건통계학전공

이환희

이환희의 석사학위논문을 인준함

2014년 6월

위 원 장	원 성 호 (인)
부위원장	성 주 헌 (인)
위 원	김 호 (인)

Abstract

Whan-hee Lee

Department of Biostatistics

Graduate School of Public Health

Seoul National University

Background: The minimum mortality temperature (MMT) is an important concept in ecology studies for explaining how temperature affects mortality. Piecewise regression and generalized additive models (GAMs) with spline methods are generally used to find the MMT; however, these methods have difficulty in estimating variance in the MMT. For instance, piecewise regression cannot reflect the nonlinear relationship between temperature and mortality, and it is computationally very intensive to estimate variance in spline methods. Therefore it is necessary to develop a new method that can tackle both problems in the MMT.

Method: We use a parametric method based on a generalized linear model (GLM). We consider nonlinear parameters to estimate MMTs and their variances are estimated by using the Delta method. The proposed method can estimate the relative risks (RRs) and their differences can be reflected by the relationship between temperature and non-accidental mortality. The proposed methodology, we use data on five Asian cities during 1992–2010 (Seoul), 1972–2010 (Tokyo, Osaka, Nagoya), and 1994–2007 (Taipei).

Result: We find that the nonlinear model detected by our methodology represents the temperature effect on mortality well. Our results show that all estimates of MMTs are located from 22–29°C and their standard errors other than that for Seoul are less than 0.4. These results are similar or more stable with those using a B-spline method and previous epidemiological studies. We also estimate RRs to detect the extreme heat effect (differences in the 90% and 99% quantile temperatures) and estimate suitable RRs and their confidence intervals.

Conclusions: Our methodology can be an useful alternative to piecewise regression or GAMs.

Keywords: MMT, GLM, Delta method, Relative risk, temperature, mortality

Student I.D.: 2013-23585

Table of Contents

그림 목차

표 목차

1. Background

The minimum mortality temperature (MMT), also called the threshold point, is a concept widely used in ecological epidemiology studies to explain the temperature at which minimum mortality is affected[1−4]. Many studies have used the MMT to estimate the relative risk (RR) of temperature changes[2, 3, 5]. In particular, research on the effects of extreme heat usually calculate RRs at 90–99% quantile temperatures based on the MMT and estimate the difference in RRs to show that extreme heat increases mortality[3, 6]. Moreover, in studies of climate change, the MMT is used to measure the adaptation value of people[7, 8]. Therefore, an exact estimation of the MMT is an important issue in environmental epidemiology studies.

There are various methods of calculating the MMT, and in particular, piecewise linear regression[9, 10] and generalized additive model (GAM)[11] with parametric and nonparametric smoothing methods[12] have often been utilized. The piecewise regression can derive a model that contains a V−shaped relationship between temperature and mortality as well as provide different RRs based on the MMT by using two temperature coefficients[9]. However, it can consider only linear effects and nonlinear relationships such as J and U shapes cannot be estimated.

To overcome this methodological limitation, recent studies used a GAM with smoothing spline methods[2, 3, 12]. However, in spite of its flexibility for nonlinear model , it cannot directly estimate variances of estimates and adopting a resampling method such as block bootstrapping[13, 14] or a Bayesian approach that considers the prior distribution of the MMT[15−17] are computationally very intensive. Furthermore These methods can lose statistical power if the sample size is small or a assumed distribution for prior is correct[13, 17].

In this report we propose new method that estimates the MMT by using a generalized linear model (GLM) [18]. The proposed approach can consider the nonlinear relationship between temperature and mortality, and variance estimation by using a parametric method overcomes the computational problems in smoothing splines. Furthermore, by nature of the proposed method, the process of estimating RRs can be easily described and their confidence intervals (CIs) can be obtained based on the asymptotic normality of their estimates. We empirically evaluate the validity of statistical testing based on the proposed method and the accuracy of CI of the difference between two estimated RRs to detect temperature effects on mortality.

The remainder of the paper is structured as follows. Section 2 briefly reviewed the basic formula of a GLM with nonlinear parameters, and introduces how to estimate parameters and calculate approximate variance in the MMT by using the Delta method. In this section, we also proposes more advanced model that can represent the complex relationships between temperature and mortality. The proposed method was utilized to model the effect of temperature on mortality for five Asian cities and results are illustrated in the Results. Last, we discuss the practical meaning of the proposed methods and provide some suggestion for further work.

2. Method

2.1 Illustrative data

We use data on daily temperature and non-accidental mortality counts for five Asian cities: Seoul (Korea, 1992–2010), Tokyo, Osaka, and Nagoya (Japan, 1972–2009), and Taipei (Taiwan, 1994–2007). The primary confounder control contains daily air pressure, relative humidity, and Fourier terms to adjust seasonality and long-term trends. The data are summarized in Table 1. We illustrate each process discussed by using R code, reproducing the results in the Appendix.

2.2 A basic model

This model is motivated by the developmental rate of *Drosophila melanogaster* example of McCullagh and Nelder[18].

2.2.1 Estimating the MMT

The basic model used to estimate the MMT with number of daily mortality counts $Y_i, i = 1,\ldots,n$, is given by

$$g(\mathrm{Y}_i) = \beta_0 + \beta_1 \mathrm{T} + \beta_2 \frac{1}{\mathrm{T} - \delta} + \sum_{k=1}^{K} \gamma_k u_{ik}.$$

To find $\widehat{\mathrm{T}}$ which satisfy $\frac{\partial \mathrm{g(Yi)}}{\partial \mathrm{T}} = 0,$

$$\hat{T} = \hat{\delta} - \sqrt{\frac{\hat{\beta}_2}{\hat{\beta}_1}} \qquad (1)$$

where $\mu \equiv \mathrm{E(Y_i)}$, g is a monotonic link function, and Y is assumed to follow exponential family distributions. β are coefficient parameters and T means daily mean temperature. The variables u_k include other predictors as covariates by the matched coefficients γ_k. All parameters are estimated by maximum likelihood estimation (MLE). Furthermore, we use a rational function of temperature and a nonlinear parameter, δ to show the nonlinear relationship between temperature and non−accidental daily mortality.

The outcome Y_i is daily count, which is assumed to originate from a quasi−Poisson distribution. This considers over−dispersion in the data with $E(Y) = \mu, Var(Y) = \sigma^2\mu$ and a canonical log link.

We estimate the MMT by finding a root for $g'(Y_i) = 0$ toward T. This equation was solved by the optim function in *R*. which is based on the Nelder–Mead algorithm[19]. The starting value of the nonlinear parameter δ is set to a maximum temperature of +2° C, and the starting values of the other parameters are set by the coefficients estimated in the GLM process, considering the starting value of δ as a constant. This starting value process can improve efficiency and calculation accuracy.

2.2.2 Variance in the MMT

From equation (2), the estimated MMT is described as fractional terms of $\hat{\beta}s$. We use the Delta method[20] to estimate variance in the MMT based on a combination of parameters, $g(\hat{\beta}, \hat{\delta})$. Delta method finds approximate variance by using Taylor series expansions of a variance function of random variables. By applying Slutsktys theorem and a normal asymptotic property based on the central limit theorem[20, 21], the Delta method can also derive variance in the function of $g(\hat{\beta}, \hat{\delta})$. Then, approximate variance of estimates for the MMT can be derived by

$$var(\hat{T}) \approx [\frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_1} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_2} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\delta}}] \cdot \hat{\Sigma} \cdot [\frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_1} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_2} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\delta}}]^T \qquad (2)$$

$$\hat{\Sigma} \geq \hat{I}(\hat{\beta}_1, \hat{\beta}_2, \hat{\delta})^{-1} \qquad (3)$$

where $\hat{\Sigma}$ is the observed fisher information matrix of the parameters. From Cramer–Rao lower bound theory, we can then use $\hat{\Sigma}$ as a minimum variance unbiased estimator of the covariance matrix[22].

2.3 An advanced model

2.3.1 Estimating the MMT

In the basic model, we used only one nonlinear parameter, δ, which applies an asymptotic line of temperature. This model assumed that the nonlinear relationship between temperature and mortality was fractional. Although the basic model is suitable for finding the MMT in normal relationships (e.g., left skewed, U−, J−, or V−shaped), if data do not follow normal cases, it becomes unsuitable. Hence, because the real−world relationship between temperature and mortality may be asymmetric, many studies use GAMs with smooth splines. However, we can still reflect the complex relationship and estimate the MMT by using an advanced GLM with additional nonlinear parameters or the transformation of variables. The equation of such an advanced model can generally be written as

$$g(Y_i) = \beta_0 + \beta_1 f(T) + \beta_2 f(T,\delta) + \beta_3 f(T,\theta) + \sum_{k=1}^{K} \gamma_k u_{ik}, \quad (4)$$

where $f(.)$ is a nonlinear function as in equation (1) and θ contains an asymptotic line of cold temperature. By adding additional nonlinear parameters, our model can thus contain two separate nonlinear relationships between temperature and mortality. In addition, $f(.)$ can be more flexible, like $1/T, e^T, log(T), \ldots etc$.

In this study, we use the advanced model in (6) to consider the cold effect. This can be written as

$$g(Y_i) = \beta_0 + \beta_1 log(T) + \beta_2 \frac{1}{T-\delta} + \beta_3 log(T+\theta) + \sum_{k=1}^{K} \gamma_k u_{ik}. \quad (5)$$

Then, we derive the MMT by using a function of the estimated parameters:

$$\hat{T} = h(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\delta}, \hat{\theta}). \quad (6)$$

In this case, $h(.)$ is a function of the cubic formula [23]. In other cases, we can find the estimated MMT by following a similar process. Under the 5th degree parametric space, we can express the MMT as a combination of coefficients and nonlinear parameters.

2.3.2 Variance in the MMT in the advanced model

As in equation (3), we can also estimate variance in the MMT ($\hat{T}$ of (7)) by using the Delta method:

$$var(\hat{T}) \approx [\frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_1} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_2} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_3} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\delta}} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\theta}}] \cdot \hat{\Sigma} \cdot [\frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_1} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_2} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\beta}_3} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\delta}} \frac{\partial \hat{\mathbf{T}}}{\partial \hat{\theta}}]^T . \quad (7)$$

Under this process, we can find variance in the MMT if we use more nonlinear parameters than those shown in formula (1). By estimating variance in the MMT, we can also conduct several statistical tests about the MMT. We also carry out hypothesis testing to validate the MMT (H_0: MMT is a specific value) by using a Wald test (Wald 1943) and a homogeneity test among the MMTs of the sampled cities by using a Cochran Q (Cochran 1954) test[24]. Many other statistical tests for MMTs are also available.

2.4 RR by using a GLM with nonlinear parameters

RR is the ratio of the probability of an event occurring in an exposed group to the probability of the event occurring in a comparison, non-exposed group[25]. $exp(\hat{\beta})$ for a Poisson log linear model becomes RR[24, 25]. However, because it can only consider a linear effect, many ecological epidemiology studies have begun to use spline methods to estimate RR, with the number of events at the MMT used as the denominator.

Not only can our methodology calculate RR, it can also capture the non-linear associations of temperature. Because our model do not use a linear predictor, the standard error of RR is hard to calculate directly. To solve this problem, we use the Delta method again.

2.4.1 RRs and CIs

The procedure for calculating the expected value of RR at x°C from the nonlinear parameters can be written as

$$\hat{RR}_x = \frac{\hat{\mu}_x}{\hat{\mu}_{MMT}} = \frac{exp(\hat{\beta}_0 + \hat{\beta}_1 f(T_x) + \hat{\beta}_2 f(T_x, \hat{\delta}) + \hat{\beta}_3 f(T_x, \hat{\theta}) + \hat{c}ovs)}{exp(\hat{\beta}_0 + \hat{\beta}_1 f(T_{MMT}) + \hat{\beta}_2 f(T_{MMT}, \hat{\delta}) + \hat{\beta}_3 f(T_{MMT}, \hat{\theta}) + \hat{c}ovs)}$$

$$= exp[\hat{\zeta}] \quad (8)$$

$$var(\hat{\zeta}) \approx [\frac{\partial\hat{\zeta}}{\partial\hat{\beta}_1} \frac{\partial\hat{\zeta}}{\partial\hat{\beta}_2} \frac{\partial\hat{\zeta}}{\partial\hat{\beta}_3} \frac{\partial\hat{\zeta}}{\partial\hat{\delta}} \frac{\partial\hat{\zeta}}{\partial\hat{\theta}}][\hat{\Sigma}][\frac{\partial\hat{\zeta}}{\partial\hat{\beta}_1} \frac{\partial\hat{\zeta}}{\partial\hat{\beta}_2} \frac{\partial\hat{\zeta}}{\partial\hat{\beta}_3} \frac{\partial\hat{\zeta}}{\partial\hat{\delta}} \frac{\partial\hat{\zeta}}{\partial\hat{\theta}}]^T .$$

Then, the 95% CI of RR is obtained from a simple equation with a standard error of ζ , $s.e(\hat{\zeta}) = \sqrt{var(\hat{\zeta})}$

$$95\%\, C.I.\, of\;\; RR = exp[\hat{\zeta} - 1.96 s.e(\hat{\zeta}), \hat{\zeta} + 1.96 s.e(\hat{\zeta})] \tag{9}$$

2.4.2 Difference in RR

The difference in RR between x°C and y°C is described as the ratio of two estimated RRs as in (9):

$$\frac{\dfrac{\hat{\mu}_x}{\hat{\mu}_{MMT}}}{\dfrac{\hat{\mu}_y}{\hat{\mu}_{MMT}}} = \frac{\hat{\mu}_x}{\hat{\mu}_y} . \tag{10}$$

Many ecological studies have used such a ratio (e.g., between exposure levels) to detect RR. However, our model can define differences in RRs by using (12); therefore, we can calculate this and its variance.

3. Results

3.1 Model selection

The analysis presented in this section is based on models (1) and (5). These models are fitted with a quasi Poisson family to control confounders. Two Fourier terms (one sine and one cosine, period=365.25) of time and a linear term of time are used to describe seasonality and long−term trends. Further, the models include indicator variables for day of the week, natural cubic splines of relative humidity and air pressure, and indicator variables for an outbreak of influenza following several epidemiology studies of time series analysis[12, 26−28].

The effect of daily mean temperature is investigated through a combination of nonlinear parameters. In this study, we fit data on each city to models (1) and (5) and then select the model and number of parameters based on modified Akaike information criteria for models with quasi likelihood for an overdispersed response variable (QAIC) [11, 29], as given by

$$QAIC = -2l(\hat{\theta}) + 2\hat{\sigma}^2 k \tag{11}$$

where l is the log likelihood of the fitted model with a quasi Poisson distribution, $\hat{\theta}$ is a parameter, and $\hat{\sigma}^2$ is the estimated overdispersion parameter. k is the number of parameters. Although we select a model between (1) and (5) based on a lower QAIC, if the difference in the QAIC is too small to add more parameters or parameters do not converge with the other starting values, we use the basic model instead. All the estimated MMTs, variances in MMTs, and the QAIC of each city are described in Table 2.

Moreover, to consider the lagged effect of temperature on mortality, we apply five types of lag structures according to previous studies [28, 30−33]: averages from 0 days to 1 day (lag 0−1), 2 days to 6 days (lag 2−6), and 7 days to 13 days (lag 7−13) as well as a moving average temperature for 1 week (MV0−6) and 2 weeks (MV0−13).

To fit the data to model (5), we take the parallel translation in the direction of the x axis because of the log domain. The degree of translation is set to minimize movements that can make the log domain non−zero. This process does not affect the estimated variance.

3.2 Results

Table 2 compares the estimated MMTs, standard errors, and CIs of the MMT from model (1) with the estimated values from the model using the B−spline method (degrees of freedom (df) = 5). The results presented in Table 2 are based on the moving average for two weeks (i.e., MV0−13) to consider the long−term effect of temperature. By using a QAIC or parameter convergence, the MMT of 3 Japan cities (Tokyo, Osaka, Nagoya) were estimated with model (5), whereas the MMTs of the other two cities are calculated from model (1). Except for Seoul, the standard errors of the estimated MMTs are shown to be suitable, with the MMTs of the five cities ranging between 22° C and 29° C.

As a sensitivity analysis, we compare the degrees of freedom of B−splines with 3 degrees and confirm that our results are consistently reasonable for both spline methods. The results of the estimated MMTs from the other lag structures (lag 0−1, lag 2−6, lag 7−13, MV0−6) and of the sensitivity analysis are described in the Appendix.

Figure 1 shows the nonlinear relationships between temperature and non-accidental mortality for the five different lags in Seoul and Tokyo. This figure shows unclear MMTs in the Seoul plot compared with Tokyo, especially for lag 0-1. For this reason, the range of the standard error of the MMT is wide in Seoul (Table 2). By contrast, in Tokyo we find relatively clear MMTs for the lags, although the MMT of lag 0-1 has relatively large variance. Table 1 describes the estimated MMTs and their standard errors for the five types of lag for the five cities.

Further, to verify the reliability of our estimated MMT, we compare our result with the B-spline method generally used in environmental research with the same confounder in our model (see Table 2). We find that the MMTs estimated by using the GAM of all cities were quite similar with the MMTs from our methodology. And in order to test reliability of out methodology, we calculated the empirical standard errors of our model and GAM by using block bootstrapping, the adjusted bootstrapping method applied to time series [13, 34, 35]. In this case, we selected 14 length blocks same with lag distance and 1000 times subsampling. We found that estimated MMTs using our methodology had smaller standard errors than GAM. Moreover estimated MMTs of our methodology showed more symmetrical and concentrated on empirical mean values of each city. Other results at different degrees of freedom also seem to be similar. Accordingly, we assume that our method is suitable compared with the GAM method mostly used in time series analysis in ecology studies.

The estimated MMTs are also close to the results of a previous epidemiology study [2], which found that the quantiles of the MMT were situated at 89% of the temperature in Korea, 86% in Japan, and 62% in Taiwan, although the confounders used in that study were slightly different. After converting these quantiles into Celsius, we find the following results: Seoul 25.12°C, Tokyo 25.70°C, Osaka 27.03°C, Nagoya 25.90°C, and Taipei 26.00°C. From these results, we conclude that our method seems to be reasonable for estimating MMTs compared with both previously used spline methods and epidemiological research.

The RRs of the effect of temperature on mortality are provided in Figure 1 with the 95% CI (dashed line). The RRs of 90% and 99% of temperature and their differences are presented in Table 3 with 95% CIs. All the estimated RRs are based on the MMT as a reference. When the difference in RR is 90% compared with 99%, almost all cities show values from 1.03 to 1.07. Table 3 also shows the RRs estimated from using the B-spline method, confirming that the RRs from our method are similar to those estimated by using the B-spline method. This finding also means that our methodology seems to be suitable for finding RRs in environmental studies compared with general frameworks.

4. Discussion

In this study, we describe the parametric estimating of the MMT and its variance by using a GLM with nonlinear parameters. The modeling framework proposed herein may be applicable to environmental studies, complementing previous methods that have failed to estimate variance in the MMT. The analysis of data on five Asian cities during 1972–2010 offers some evidence for the potential of our methodology for statistically estimating the MMT and its variance when temperature and mortality have a nonlinear relationship, which precludes the use of spline methods or piecewise regression. We believe this methodological framework thus represents a useful tool to find the temperature effects in environmental studies and other ecology research.

Although our model has advantages over other methods, however, specific issues arise when it is applied to common models to estimate the MMT and RR. These are discussed below.

Simple procedure. We can derive an MMT by using a combination of parameters, which simplifies our model process for deriving variance in the MMT and estimating RRs, their differences, and CIs compared with spline methods. This way of estimating the MMT is also applied in polynomial regression, a form of our advanced model. Although it is hard to calculate when the equation of the MMT has a high degree (over 4th), this does not prevent the estimation of an MMT. Indeed, the results from the basic models confirm that the relationships between temperature and mortality near MMTs are U- or J-shaped, whereas more complex relationships are shown at cold temperatures. Further, according to previous epidemiology studies[2, 3], an MMT should be located above the mean temperature of each city. Our results similarly show that the MMTs of cities are located above the mean temperature. Therefore, we conclude that our simple models (1) and (6) are sufficient for estimating a reasonable MMT and its variance.

Sample size. Unlike spline methods, our parametric method does not suffer problems with regard to sample size. In time series analyses of environmental studies, sample size is not commonly a problem because they use data collected over a minimum of a few years. However, some studies with small sample sizes (e.g. clinical trials), which employ a cohort study to find temperature effects, find it difficult to use the GAM with spline methods. In addition, when the study period is short, our method is advantageous for estimating the MMT and its variance.

Nonlinear parameters. Although our method involves only two nonlinear parameters, more nonlinear parameters can be added if the researcher desires. However, if a model has too many nonlinear parameters, it not only makes estimating the MMT more complex, but also erodes the goodness of fit.

Therefore, employing a deviance test or comparing information criteria is required to select fittable nonlinear parameters. We thus recommend using three or less nonlinear parameters in one model.

Likelihood process. We use MLE to estimate the parameters (including nonlinear ones). Parameter variance is occasionally found to be negative definite when we use too complex a nonlinear function or set an inappropriate starting value. Although an approximated estimation using a Taylor expansion is presented as a one of the solutions to this problem, this is not a suitable method because it fails to show consistent starting values and shows a larger information criterion value than that under our method. Moreover, likelihood estimation with nonlinear parameters may not provide suitable results because the shape of the likelihood function is not fittable to convergence. In rational form, the shape of the likelihood function shows an asymmetric shape (Appendix Figure7) because of the defined domain. To address these likelihood problems, we repeated the procedure to set a starting value and select a suitable function.

Stability and Flexibility. First we tested goodness of fit using the R^2 statistics. Appendix Table A1 showed that the R^2s of our methodology were quite similar with GAM (df=5). At Table2 and Figure2, we could find our methodology had stability to estimate MMTs compared with GAM. We found that the densities of MMT using our methodology were more stabilize than GAM(df=5)'s. As our conjecture, when sample size became smaller because of block bootstrapping and degrees of freedom became bigger with smoothing, estimation of GAM was unstable[36]. However *o*ur methodology has less flexibility than GAM methods because our model is based on a parametric inference. Although a parametric process has a strong advantage if data are well fitted with a parametric relation, it shows unsuitable performance compared with spline methods when data are not appropriate for use with a parametric relationship. Therefore, we must check that temperature and mortality have a parametric relationship (e.g., by drawing scatterplots).

Lagged effect. In this study, we consider the lagged effect by using five types of single lags and moving averages. Although we try to reflect a delayed effect by using a distributed lag model[37], the existence of nonlinear parameters prevents us from using such a model. However, a distributed lag model could be used in certain circumstances by making the estimated nonlinear parameter a constant value. Moreover, because formulas such as (1) or (6) are basic functions, we expect that our method could be fused and supplemented with a distributed lag nonlinear model[27].

Finally, our result uses a quasi Poisson with log link. Nevertheless, frameworks based on equation (4) and the Delta method could have general applicability (e.g., for time series data with other outcome distributions or links).

In particular, in cases of overdispersion, negative binomial or gamma likelihood could be used with our estimation process[18, 38]. More importantly, the main idea of our study is considerably general, and thus it can easily be applicable to other environmental or toxicology studies that wish to find the MMT by using regression models.

Table 1: Descriptive statistics for the study variables in the five Asian cities: Seoul (Korea, 1992–2010), Tokyo, Osaka, and Nagoya (Japan, 1972–2009), and Taipei (Taiwan, 1994–2007)

	Seoul	Tokyo	Osaka	Nagoya	Taipei
Mean Temperature(℃)	12.76	16.15	16.75	15.65	23.27
Min Temperature(℃)	−11.2	0.2	−0.8	−2.7	8.6
Max Temperature(℃)	32.75	32.40	32.50	32.40	32.10
Mean Relative Humidity(%)	62.88	62.32	63.81	67.01	76.08
Mean Press(mph)	1016.19	1009.56	1007.36	1008.00	1012.32

Table 2: The estimated MMT, its standard error (S.E.), 95% CI, and MMT estimated by using the spline method based on the moving average temperature from lag 0 to lag 13.

	MMT (basic)	s.e.	MMT (advance)	s.e.	MMT (df=5)	s.e. (empirical)	MMT (df=11)	s.e. (empirical)
Seoul	22.15	1.78	NaN	-	24.4	4.84	23.1	10.9
Tokyo	25.21	0.17	24.89	0.32	25.0	9.68	34.9	11.3
Osaka	26.50	0.28	25.84	0.56	26.0	4.09	24.8	11.1
Nagoya	25.00	0.32	22.16	0.32	24.9	8.15	23.7	10.7
Taipei	28.69	0.21	NaN	-	27.5	2.89	26.6	10.4

Table 3: RR values at 90% and 99% of temperature on mortality with the MMT as a reference value and its difference. All numbers are based on the moving average temperature from lag 0 to lag 13. Results of Seoul and Taipei were estimated by the basic model (1), other results were from the advanced model (5).

	90%RR		99%RR		RR difference		RR difference (GAM, df=5)
Seoul	1.003	(0.998,1.007)	1.02	(1.0008,1.033)	1.018	(1.010,1.026)	1.071
Tokyo	1.004	(1.002,1.006)	1.054	(1.044,1.063)	1.022	(1.018,1.027)	1.06
Osaka	1.006	(1.000,1.011)	1.04	(1.025,1.1055)	1.034	(1.022,1.046)	1.043
Nagoya	1.008	(1.000,1.017)	1.063	(1.043,1.082)	1.027	(1.018,1.035)	1.079
Taipei	1.003	(1.000,1.007)	1.04	(1.023,1.057)	1.037	(1.023,1.051)	1.014

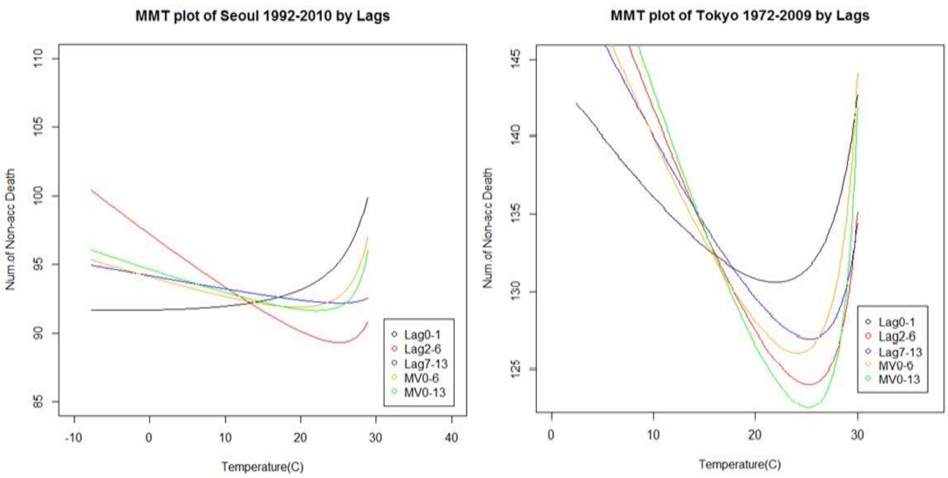

Figure 1: Temperature−Non accidental death plots of Seoul and Tokyo by lag. Regression plots based on the basic model by using the five types of lagged relationships. The MMTs are not clearly shown in the plot of Seoul (left), while those of Tokyo are clear (about 25–26° C).

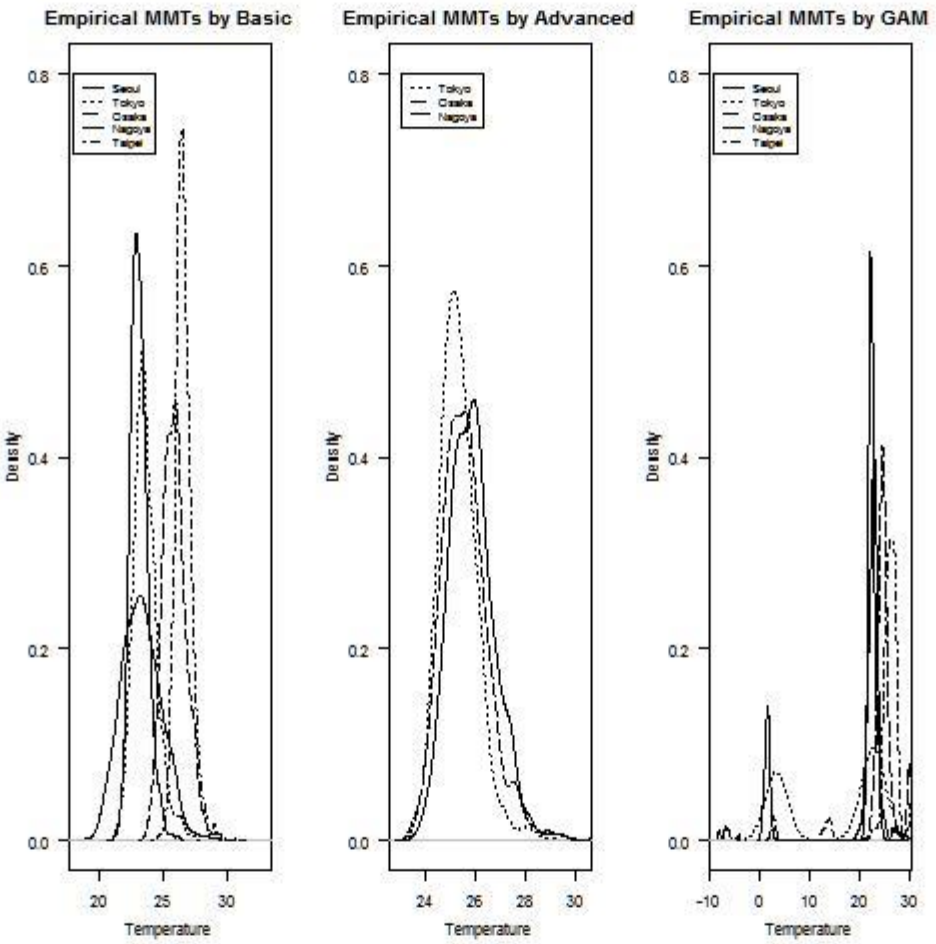

Figure 2 : Density plots of empirical MMTs using block bootstrap. Distribution of empirical MMTs for each cities using block bootstrap (block lengths=14, the number of bootstrap required=1000) with moving averaged for 2 weeks. Empirical MMTs using the basic model or advanced model were more symmetrical and centralized than the generalized additive model (GAM, B-spline method with 5 df) models'.

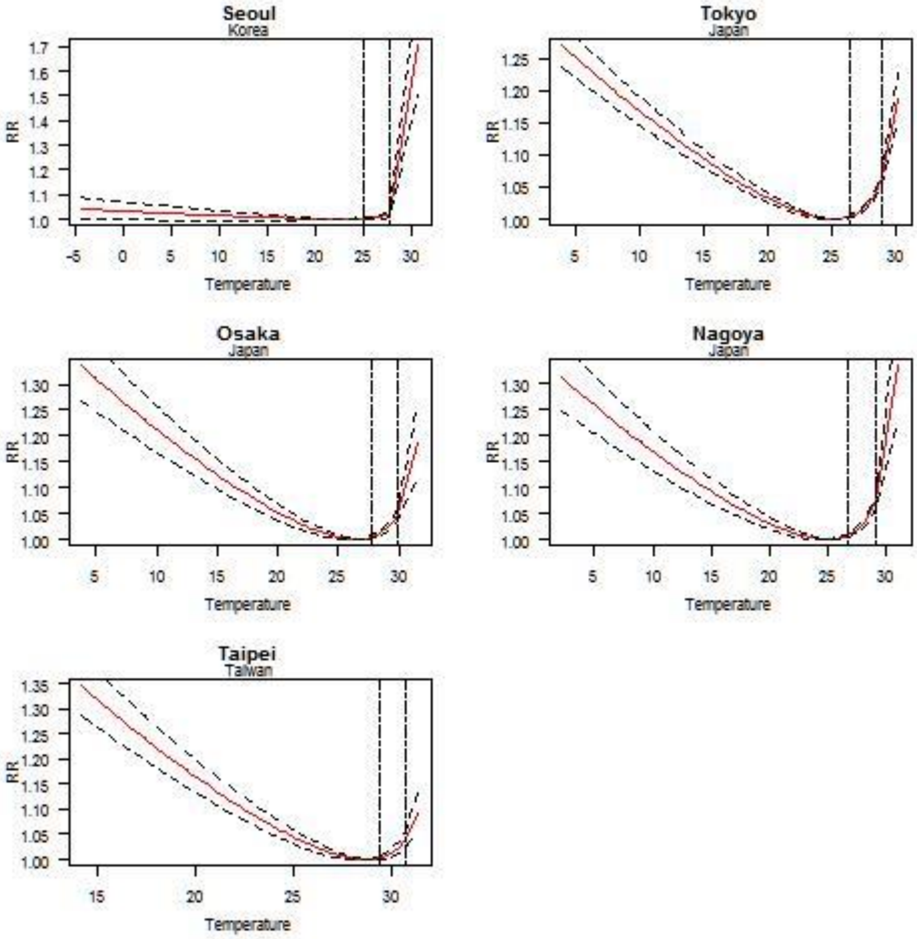

Figure 3 : RR plots for the five Asian cities. RR plots for the five Asian cities based on the basic model, moving average mean temperature for 2 weeks. The straight line represents the estimated RRs in each city and the dashed line is their 95% CIs. The vertical line shows the 90% and 95% (from left) quantiles of temperature for each city. All plots except Seoul show that the estimated MMT of each city is lower than 90% of temperature.

.

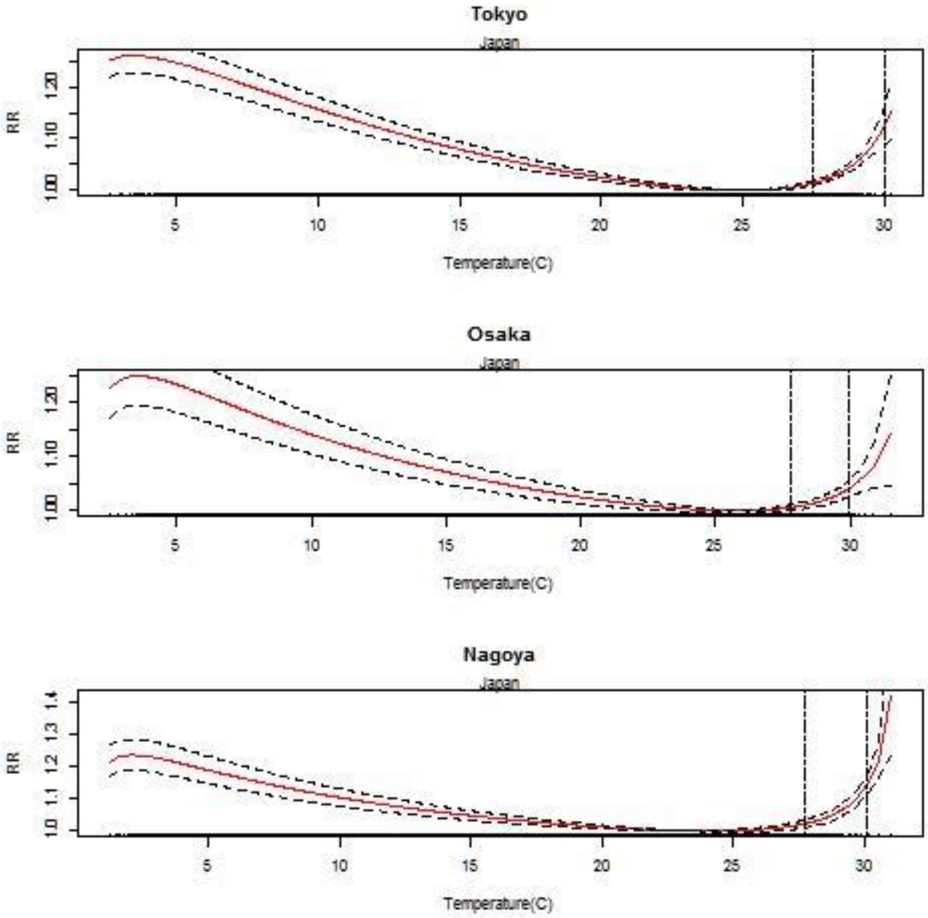

Figure 4: RR plots of 3 Japan cities estimated by using the advanced model. The RR plots of 3 Japan cities estimated by using the advanced model based on moving average mean temperature for 2 weeks. The straight line represents the estimated RRs in each city and the dashed line is their 95% CIs. The vertical line shows the 90% and 95% (from left) quantiles of temperature for each city. Compared with the plot of Tokyo in Figure 2, the cold effect (left part of the plot) is slightly reflected here.

Appendix 1. Figures and Tables

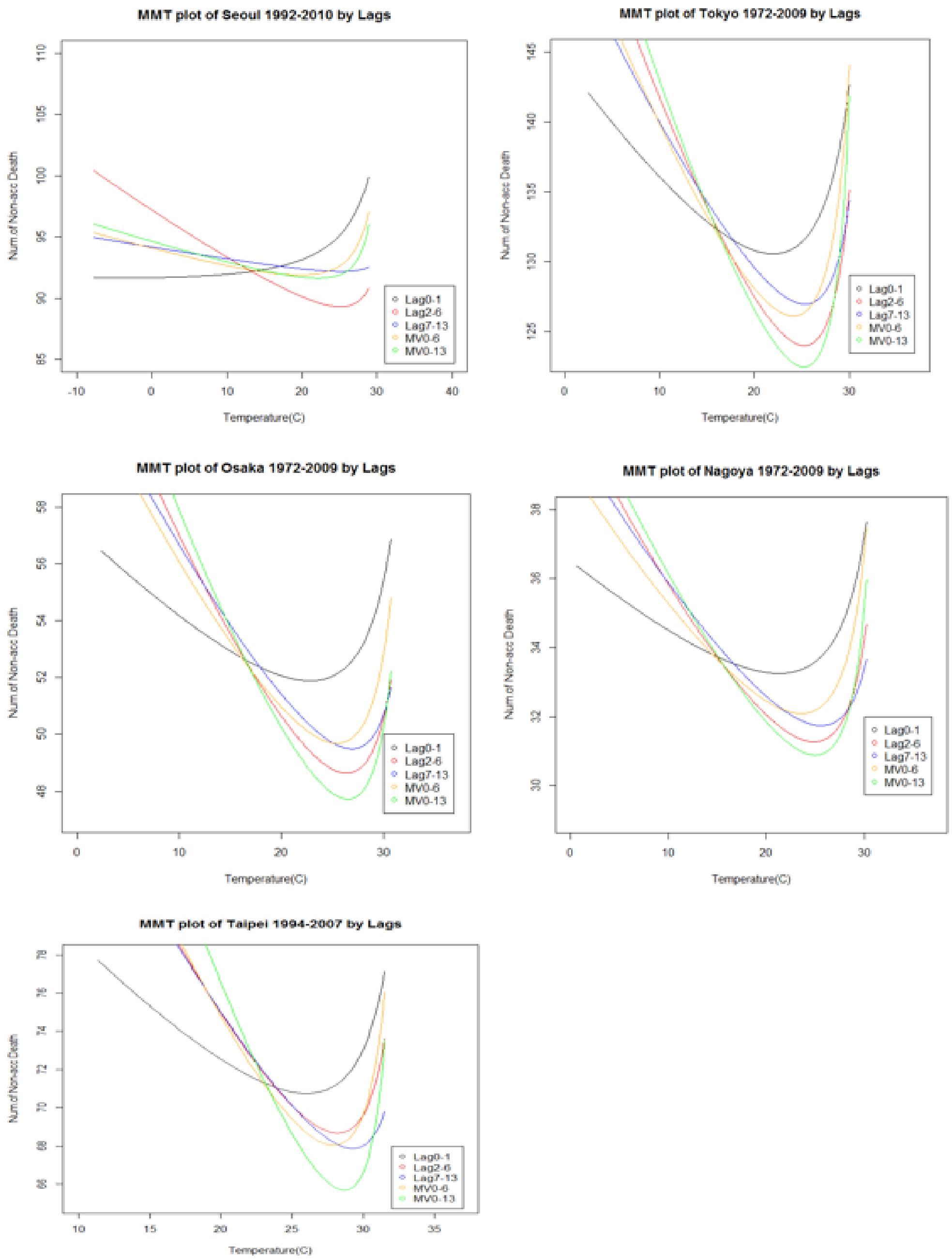

Figure A 1: Plots to find MMTs with a basic model by type of lags and by 5 cities. From our methodology framework, there are clear MMTs between 20C−30C in all cities except for Seoul. And in all 5 cities, MMTs of Lag0−1 (average of lag0 and lag1) seemed less obviously than other lags types.

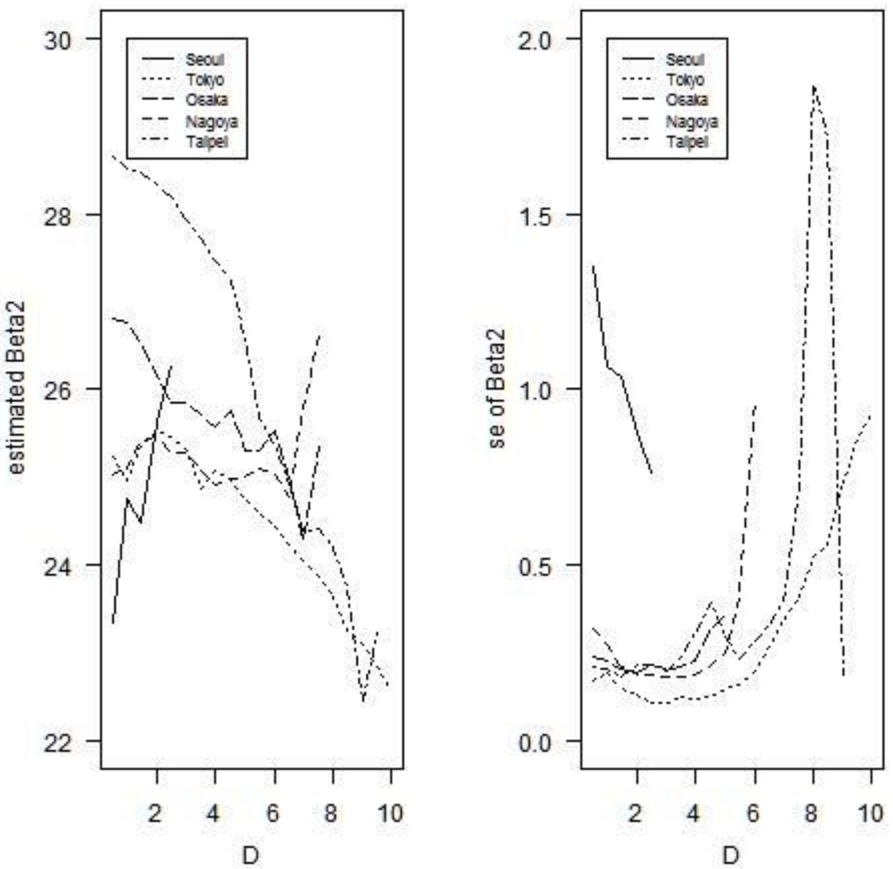

Figure A 2: Change of estimated β2 and its standard errors of basic model. We estimated β_2 based on changing temperature (original temperature−0.5*D). The more inflection of temperature disappears, the smaller β_2 and larger standard errors were estimated.

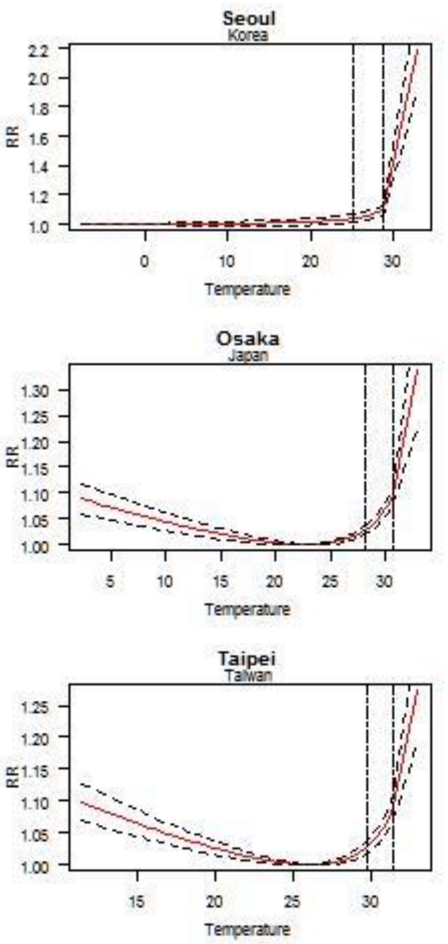

Figure A 3: Relative Risk(RR) plots for the 5 Asian cities, based on the averaged temperature of lag0 and lag1. Dashed vertical lines mean 90, 99% quantile of temperatures.

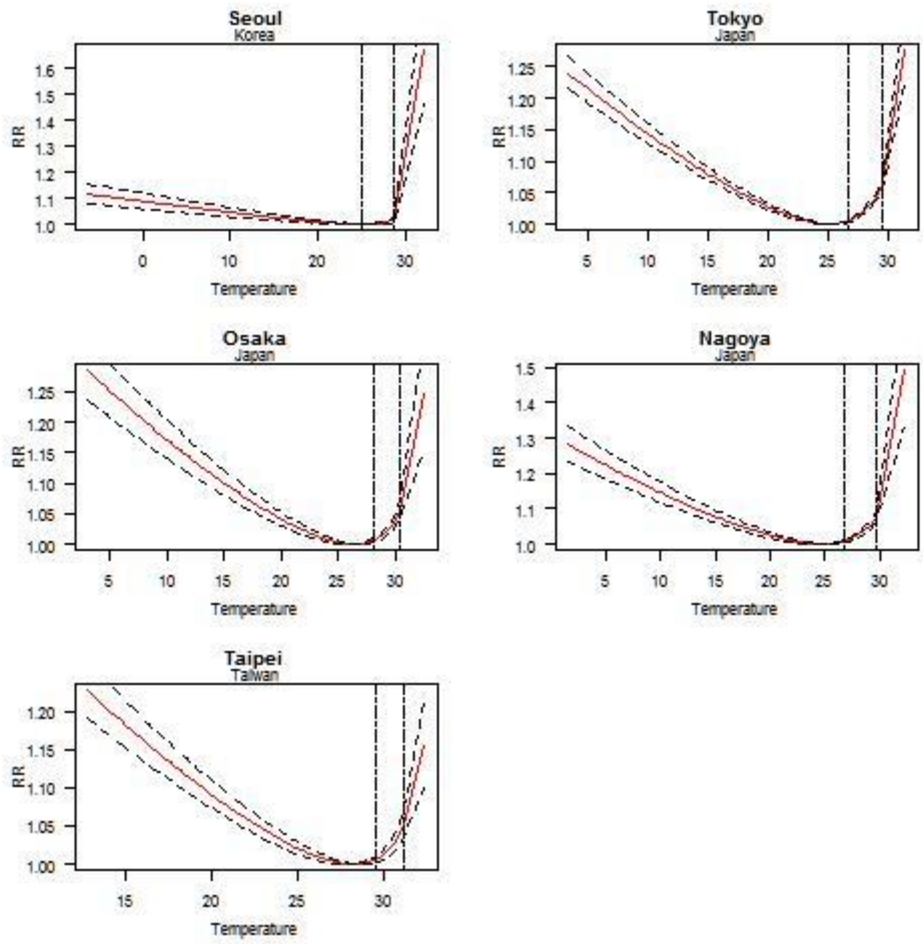

Figure A 4: Relative Risk(RR) plots for the 5 Asian cities, based on the averaged **temperature from lag2 and lag6.** Dashed vertical lines mean 90, 99% quantile of temperatures.

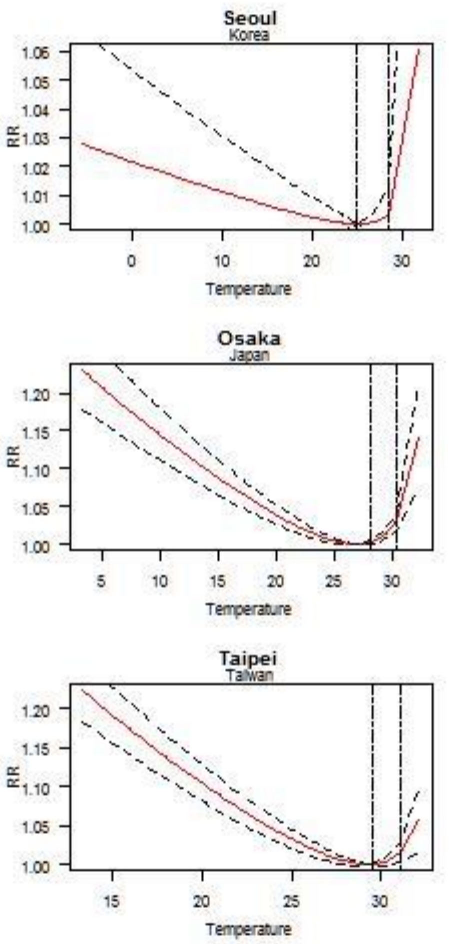

Figure A 5: Relative Risk(RR) plots for the 5 Asian cities, based on the averaged temperature from lag7 and lag13. Dashed vertical lines mean 90, 99% quantile of temperatures.

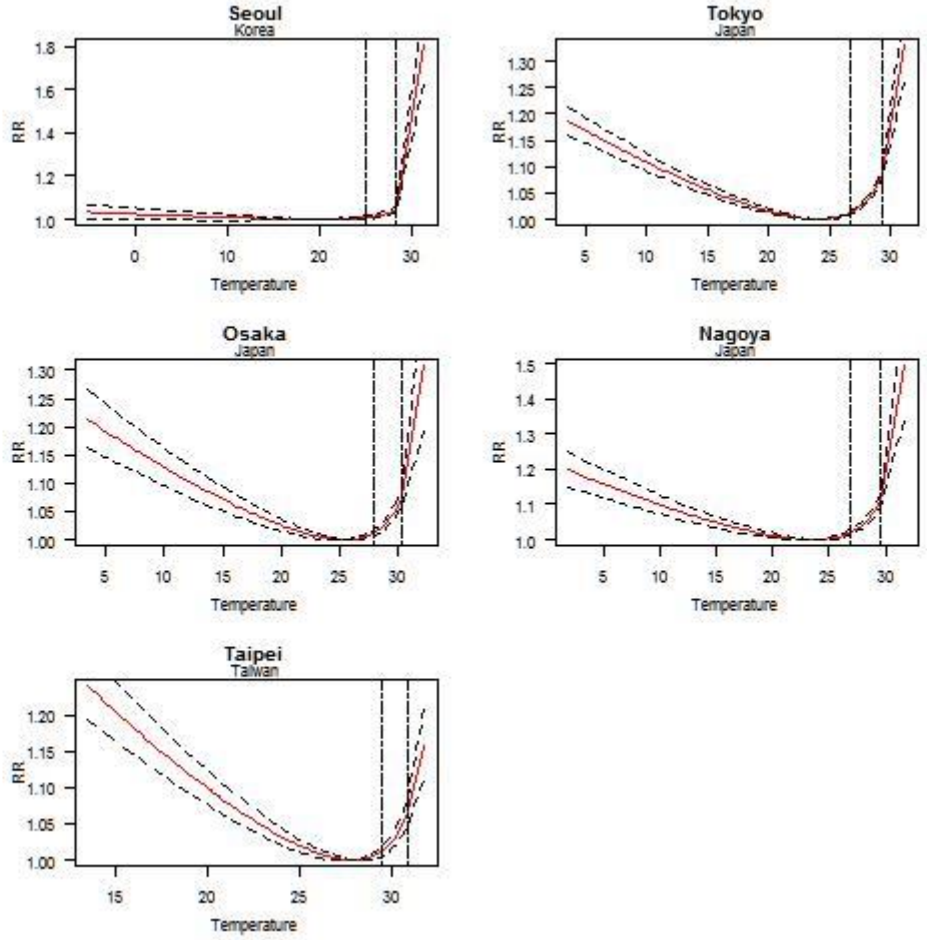

Figure A 6: Relative Risk(RR) plots for the 5 Asian cities, based on the moving average temperature for 1 week. Dashed vertical lines mean 90, 99% quantile of temperatures.

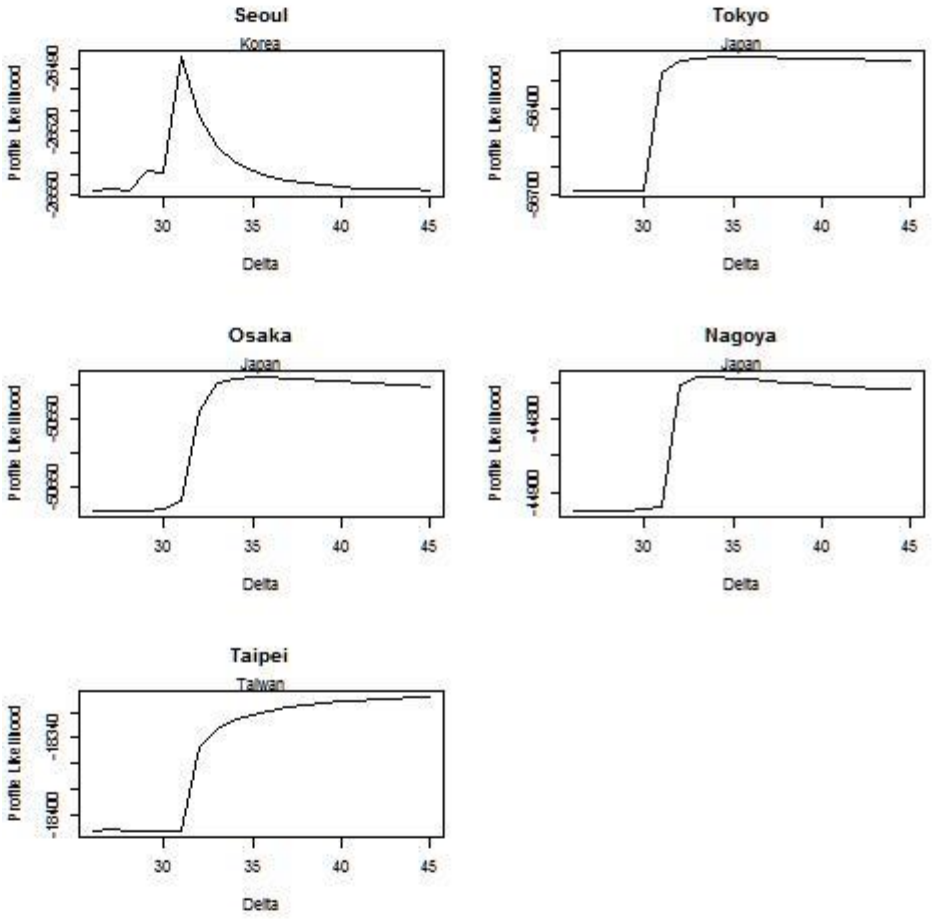

Figure A 7: **Profile likelihood plots of delta (δ) in formula (1) by 5 cities.** Shapes of profile likelihood of 5 cities were not symmetric.

Table A 1. R square statistics by models. We found goodness of fit of our methodology was similar with generalized additive model (GAM).

	Basic Model	Advanced Model	GAM (df=5)
Seoul	0.211	0.563	0.223
Tokyo	0.757	0.563	0.757
Osaka	0.516	0.558	0.516
Nagoya	0.563	0.563	0.563
Taipei	0.463	0.558	0.469

Table A 2. The estimated MMT, and its standard error (S.E.). MMT estimated by our basic model (1) and the spline method (df=4,5,11) based on the average temperature from lag 0 to lag 1.

	model(1)	**se**	**df=4**	**df=5**	**df=11**
Seoul	-4.137	24.992	-3.1	-15.7	-15.7
Tokyo	21.977	0.334	19.6	14.8	17.4
Osaka	22.799	0.662	21.8	22.4	23.2
Nagoya	21.295	0.657	21.1	23.4	-2.8
Taipei	26.035	0.442	23.6	23.7	23

Table A 3 The estimated MMT, and its standard error (S.E.). MMT estimated by our basic model (1) and the spline method (df=4,5,11) based on the average temperature from lag 2 to lag 6.

	model(1)	**se**	**df=4**	**df=5**	**df=11**
Seoul	25.119	0.745	24.2	24.8	23.7
Tokyo	25.266	0.169	23.9	25.1	24.6
Osaka	26.382	0.287	25.5	25.6	24.6
Nagoya	24.864	0.317	24	24.8	23.7

Taipei	28.213	0.265	25.5	25.7	26.4

Table A 4 The estimated MMT, and its standard error (S.E.). MMT estimated by our basic model (1) and the spline method (df=4,5,11) based on the average temperature from lag 7 to lag 13

	model(1)	**se**	**df=4**	**df=5**	**df=11**
Seoul	25.155	3.188	24.5	24.9	23.5
Tokyo	25.343	0.254	23	25.2	26.1
Osaka	26.900	0.353	25.8	26	24.1
Nagoya	25.591	0.386	24.5	25	22.8
Taipei	29.268	0.357	32	32	30.9

Table A 5 The estimated MMT, and its standard error (S.E.). MMT estimated by our basic model (1) and the spline method (df=4,5,11) based on the moving average temperature for 1 weeks

	model(1)	**se**	**df=4**	**df=5**	**df=11**
Seoul	19.683	2.172	22.7	24.1	22.8
Tokyo	24.119	0.232	22.4	24.4	24.2
Osaka	25.408	0.383	24.4	25	24.2
Nagoya	23.565	0.450	23	24.3	23.5
Taipei	27.906	0.274	25.3	25.4	22.9

Table A 6 The estimated MMT, and its standard error (S.E.). MMT estimated by our basic model (1) and the spline method (df=4,5,11) based on the moving average temperature for 2 weeks

	model(1)	**se**	**df=4**	**df=5**	**df=11**
Seoul	22.149	1.784	23.4	24.4	23.1
Tokyo	25.210	0.173	23.9	25	24.9

Osaka	26.501	0.277	25.8	26	24.8
Nagoya	25.002	0.324	24.1	24.9	23.7
Taipei	28.680	0.212	27	27.5	26.6

Table A 7 RR values at 90% and 99% of temperature on mortality with the MMT as a reference value and its difference. Numbers in parentheses represent the 95% CI based on the average temperature from lag 0 to lag 1.

	MMT	RR of 90%temp	LCL	UCL	RR of 99% temp	LCL	UCL	RR-diff	LCL	UCL
model (1)										
Seoul	-4.14	1.040	1.012	1.069	1.092	1.058	1.128	1.050	1.040	1.060
Tokyo	21.98	1.022	1.018	1.027	1.093	1.083	1.102	1.069	1.063	1.074
Osaka	22.80	1.028	1.019	1.037	1.096	1.078	1.113	1.066	1.055	1.077
Nagoya	21.30	1.030	1.021	1.040	1.132	1.112	1.152	1.098	1.087	1.110
Taipei	26.04	1.028	1.018	1.039	1.090	1.069	1.112	1.060	1.048	1.073
splines (df=5)										
Seoul	24.28	1.006	1.004	1.008	1.115	1.099	1.132	1.108		
Tokyo	-0.60	1.018	0.984	1.054	1.090	1.052	1.130	1.071		
Osaka	25.90	1.025	1.021	1.030	1.085	1.066	1.103	1.058		
Nagoya	-2.90	1.007	0.943	1.076	1.125	1.050	1.206	1.117		
Taipei	27.10	1.037	1.027	1.048	1.079	1.057	1.101	1.040		

Table A 8 RR values at 90% and 99% of temperature on mortality with the MMT as a reference value and its difference. Numbers in parentheses represent the 95% CI based on the average temperature from lag 2 to lag 6.

lag26	**MMT**	**RR of 90%temp**	**LCL**	**UCL**	**RR of 99% temp**	**LCL**	**UCL**	**RR-diff**	**LCL**	**UCL**
model(1)										
Seoul	25.12	1.000	1.000	1.000	1.014	1.004	1.024	1.014	1.004	1.024
Tokyo	25.27	1.005	1.003	1.007	1.063	1.055	1.071	1.058	1.051	1.065
Osaka	26.38	1.006	1.002	1.009	1.054	1.041	1.068	1.048	1.037	1.059
Nagoya	24.86	1.008	1.004	1.013	1.076	1.062	1.091	1.068	1.057	1.078
Taipei	28.21	1.007	1.002	1.012	1.051	1.034	1.068	1.044	1.031	1.057
splines (df=5)										
Seoul	24.95	1.000	1.000	1.000	1.077	1.062	1.092	1.077		
Tokyo	0.66	0.869	0.839	0.901	0.920	0.886	0.956	1.059		
Osaka	25.72	1.012	1.006	1.017	1.052	1.034	1.071	1.040		
Nagoya	24.04	1.012	1.006	1.019	1.098	1.079	1.117	1.084		
Taipei	27.52	1.017	1.008	1.025	1.030	1.008	1.052	1.013		

Table A 9 RR values at 90% and 99% of temperature on mortality with the MMT as a reference value and its difference. Numbers in parentheses represent the 95% CI based on the average temperature from lag 7 to lag 13.

lag713	**MMT**	**RR of 90%temp**	**LCL**	**UCL**	**RR of 99% temp**	**LCL**	**UCL**	**RR-diff**	**LCL**	**UCL**

model(1)										
Seoul	25.16	1.000	1.000	1.000	1.003	0.993	1.013	1.003	0.993	1.013
Tokyo	25.34	1.003	1.001	1.004	1.037	1.029	1.045	1.034	1.027	1.041
Osaka	26.90	1.002	1.000	1.004	1.031	1.019	1.044	1.029	1.019	1.040
Nagoya	25.59	1.002	1.000	1.005	1.039	1.026	1.052	1.036	1.025	1.047
Taipei	29.27	1.000	0.999	1.001	1.015	1.002	1.029	1.015	1.003	1.027
splines (df=5)										
Seoul	25.34	1.000	0.999	1.000	1.046	1.031	1.060	1.029		
Tokyo	1.07	0.933	0.900	0.968	0.958	0.922	0.995	0.943		
Osaka	26.00	1.005	1.000	1.010	1.026	1.009	1.044	1.010		
Nagoya	-0.60	0.812	0.763	0.864	0.839	0.786	0.895	0.825		
Taipei	27.37	1.001	0.992	1.010	0.981	0.959	1.003	0.965		

Table A 10 **RR values at 90% and 99% of temperature on mortality with the MMT as a reference value and its difference.** Numbers in parentheses represent the 95% CI based on the moving average temperature for 2 weeks.

MV07	**MMT**	**RR of 90%temp**	**LCL**	**UCL**	**RR of 99% temp**	**LCL**	**UCL**	**RR-diff**	**LCL**	**UCL**
model(1)										
Seoul	19.68	1.008	1.000	1.016	1.041	1.024	1.058	1.033	1.023	1.043
Tokyo	24.12	1.012	1.009	1.015	1.088	1.079	1.098	1.076	1.068	1.083
Osaka	25.41	1.012	1.007	1.018	1.074	1.059	1.089	1.061	1.050	1.072
Nagoya	23.56	1.019	1.011	1.027	1.106	1.089	1.123	1.085	1.074	1.096
Taipei	27.91	1.012	1.005	1.018	1.067	1.048	1.086	1.055	1.041	1.068
splines (df=5)										
Seoul	24.73	1.002	1.001	1.002	1.100	1.085	1.115	1.082		
Tokyo	1.38	0.919	0.887	0.951	0.990	0.954	1.027	0.974		
Osaka	25.20	1.020	1.014	1.027	1.079	1.060	1.098	1.061		
Nagoya	24.58	1.024	1.018	1.029	1.131	1.111	1.151	1.112		
Taipei	27.73	1.019	1.011	1.027	1.051	1.028	1.074	1.034		

Appendix 2. R code

```
library(TTR)
library(MASS)
library(tsModel)
library(splines)
library(dlnm)

############## basic model ###############

total<-read.csv("E:\\ms journal\\master\\totaldata.csv")   ###5 Asia data


# seoul=11,tokyo=1131 osaka=1271 nagoya=1231 taipei=1

total$newcode[total$citycode == 1] <- 5
total$newcode[total$citycode == 11] <- 1
total$newcode[total$citycode == 1131] <- 2
total$newcode[total$citycode == 1271] <- 3
total$newcode[total$citycode == 1231] <- 4

### MMT results matrix

MMTmat<-matrix(NA,5,5)
colnames(MMTmat)<-c("MMT","se","df=4","df=5","df=11")
rownames(MMTmat)<-c("seoul","tokyo","osaka","nagoya","taipei")

### RR results

RRresult<-matrix(NA,5,10)
rownames(RRresult)<-c("Seoul","Tokyo","Osaka","Nagoya","Taipei")
colnames(RRresult)<-c("MMT","RR of 90%temp","LCL","UCL","RR of 99%
temp","LCL","UCL","RR-diff","LCL","UCL")

##### fit the basic model

for(i in 1:5){

total<-read.csv("E:\\ms journal\\master\\totaldata.csv")
total$newcode[total$citycode == 1] <- 5
total$newcode[total$citycode == 11] <- 1
total$newcode[total$citycode == 1131] <- 2
total$newcode[total$citycode == 1271] <- 3
total$newcode[total$citycode == 1231] <- 4

total<-subset(total,newcode==i)
temp<-total$meantemp
```

```
nonacc_tot<-total$nonacc_tot
humid<-total$meanhumi
press<-total$meanpress
influ<-total$influenza_new
total$time2<-seq(1,1000,len=nrow(total))
time2<-total$time2
dow<-cbind(total$dowT,total$dowW,total$dowTh,total$dowF,total$dowS,total$dowSu)
total$time<-seq(1,nrow(total))
time<-total$time

#time trend adjustment
fourier<-harmonic(time,nfreq=1,period=365.25)

#gam for humid and press
nsHumid<-ns(humid,df=3)
nsPress<-ns(press,df=3)

#select lag structures
#temp<-Lag(temp,1) ###lag1
#temp<-runMean(temp,2:6)      ###lag2-6
#temp<-runMean(temp,7:13)     ###lag7:13
#temp<-EMA(temp,n=7)      ###MV 0-7
temp<-EMA(temp,n=14)     ###MV 0-14


fnQ<-function(theta){       ##### make the quasi likelihood function
sum(
nonacc_tot*(theta[1]+temp*theta[2]+(1/(temp-theta[4]))*theta[3]
+fourier[,1]*theta[5]+fourier[,2]*theta[6]+dow[,1]*theta[7]+dow[,2]*theta[8]+do
w[,3]*theta[9]+dow[,4]*theta[10]
+dow[,5]*theta[11]+dow[,6]*theta[12]+nsHumid[,1]*theta[13]+nsHumid[,2]*theta
[14]+nsHumid[,3]*theta[15]
+nsPress[,1]*theta[16]+nsPress[,2]*theta[17]+nsPress[,3]*theta[18]
+influ*theta[19]+time2*theta[20])

-exp(theta[1]+temp*theta[2]+(1/(temp-theta[4]))*theta[3]
+fourier[,1]*theta[5]+fourier[,2]*theta[6]+dow[,1]*theta[7]+dow[,2]*theta[8]+do
w[,3]*theta[9]+dow[,4]*theta[10]
+dow[,5]*theta[11]+dow[,6]*theta[12]+nsHumid[,1]*theta[13]+nsHumid[,2]*theta
[14]+nsHumid[,3]*theta[15]
+nsPress[,1]*theta[16]+nsPress[,2]*theta[17]+nsPress[,3]*theta[18]
+influ*theta[19]+time2*theta[20])
,na.rm=TRUE)}

startdelta<-max(temp,na.rm=TRUE)+2     ###set starting point

fitglm<-summary(glm(nonacc_tot~temp+I(1/(temp-startdelta))
```

```
+fourier+dow+nsHumid+nsPress+influ+time2,family=quasipoisson(link=log)))

startbeta0<-fitglm$coefficients[1]
startbeta1<-fitglm$coefficients[2]
startbeta2<-fitglm$coefficients[3]
startbeta3<-fitglm$coefficients[4]
startbeta4<-fitglm$coefficients[5]
startbeta5<-fitglm$coefficients[6]
startbeta6<-fitglm$coefficients[7]
startbeta7<-fitglm$coefficients[8]
startbeta8<-fitglm$coefficients[9]
startbeta9<-fitglm$coefficients[10]
startbeta10<-fitglm$coefficients[11]
startbeta11<-fitglm$coefficients[12]
startbeta12<-fitglm$coefficients[13]
startbeta13<-fitglm$coefficients[14]
startbeta14<-fitglm$coefficients[15]
startbeta15<-fitglm$coefficients[16]
startbeta16<-fitglm$coefficients[17]
startbeta17<-fitglm$coefficients[18]
startbeta18<-fitglm$coefficients[19]

###derive estimated parameters and Hessian Matrix

fit<-optim(theta<-
c(startbeta0,startbeta1,startbeta2,startdelta,startbeta3,startbeta4,startbeta5,
startbeta6,startbeta7,
startbeta8,startbeta9,startbeta10,startbeta11,
startbeta12,startbeta13,startbeta14,startbeta15,startbeta16,startbeta17,startbeta18)
,fnQ,hessian=TRUE,control=list(fnscale=-1,maxit=10000))

### result
resultt<-cbind(fit$par[c(2:4)],sqrt(diag(    solve(-fit$hessian)[c(2:4),c(2:4)] )))
resultt


### dispersion parameter

muhat<-exp(fit$par[1]+fit$par[2]*temp+fit$par[3]*(1/(temp-fit$par[4]))
+fit$par[5]*fourier[,1]+fit$par[6]*fourier[,2]+fit$par[7]*dow[,1]+fit$par[8]*dow[,
2]
+fit$par[9]*dow[,3]+fit$par[10]*dow[,4]+fit$par[11]*dow[,5]+fit$par[12]*dow[,6]
+nsHumid[,1]*fit$par[13]+nsHumid[,2]*fit$par[14]+nsHumid[,3]*fit$par[15]
+nsPress[,1]*fit$par[16]+nsPress[,2]*fit$par[17]+nsPress[,3]*fit$par[18]
+influ*fit$par[19]+time2*fit$par[20])

pearsonresi<-(nonacc_tot-muhat)^2/(muhat)
```

```
disper<-sum(pearsonresi,na.rm=TRUE)/(length(temp)-length(fit$par))
Qvcov<-disper*  solve(-fit$hessian)[c(2:4),c(2:4)]    ### covariance matrix from
fisher Information

beta1<-fit$par[2]
beta2<-fit$par[3]
delta<-fit$par[4]

MMT<-delta-sqrt(beta2/beta1)
MMTmat[i,1]<-MMT    ### MMT


###delta method
fb1<-D(expression(delta-sqrt(beta2/beta1)),'beta1')   ###derivative
fb2<-D(expression(delta-sqrt(beta2/beta1)),'beta2')
fdel<-D(expression(delta-sqrt(beta2/beta1)),'delta')

mat<-rbind(eval(fb1),eval(fb2),eval(fdel))
mat2<-as.matrix(mat)
se2<-sqrt(t(mat2)%*%Qvcov%*%mat2)
MMTmat[i,2]<-se2

### GAM method
temp1<-onebasis(temp,fun="bs",degree=3,df=4)
model2<-
glm(nonacc_tot~temp1+fourier+time2+nsHumid+nsPress+dow+influ,family=quasipo
isson)
pred<-crosspred(temp1,model2,by=0.1)
MMTnon<-pred$predvar[which.min(pred$allRRfit)]
MMTmat[i,3]<-MMTnon

temp1<-onebasis(temp,fun="bs",degree=3,df=5)
model2<-
glm(nonacc_tot~temp1+fourier+time2+nsHumid+nsPress+dow+influ,family=quasipo
isson)
pred<-crosspred(temp1,model2,by=0.1)
MMTnon<-pred$predvar[which.min(pred$allRRfit)]
MMTmat[i,4]<-MMTnon

temp1<-onebasis(temp,fun="bs",degree=3,df=11)
model2<-
glm(nonacc_tot~temp1+fourier+time2+nsHumid+nsPress+dow+influ,family=quasipo
isson)
pred<-crosspred(temp1,model2,by=0.1)
MMTnon<-pred$predvar[which.min(pred$allRRfit)]
MMTmat[i,5]<-MMTnon
```

```
### calculating RR
#95% RR
tx<-quantile(temp,c(0.90),na.rm=TRUE)
fb1<-D(expression(beta1*(tx-MMT)+beta2*(MMT-tx)/(tx*MMT-
(tx+MMT)*delta+delta^2)),'beta1')   ###derivative
fb2<-D(expression(beta1*(tx-MMT)+beta2*(MMT-tx)/(tx*MMT-
(tx+MMT)*delta+delta^2)),'beta2')
fdel<-D(expression(beta1*(tx-MMT)+beta2*(MMT-tx)/(tx*MMT-
(tx+MMT)*delta+delta^2)),'delta')

zeta<-beta1*(tx-MMT)+beta2*(MMT-tx)/(tx*MMT-(tx+MMT)*delta+delta^2)
mat<-rbind(eval(fb1),eval(fb2),eval(fdel))
mat2<-as.matrix(mat)
se<-sqrt(t(mat2)%*%Qvcov%*%mat2) ###se of RR by delta method

#99% RR
tx2<-quantile(temp,c(0.99),na.rm=TRUE)
fb1<-D(expression(beta1*(tx2-MMT)+beta2*(MMT-tx2)/(tx2*MMT-
(tx2+MMT)*delta+delta^2)),'beta1')   ###derivative
fb2<-D(expression(beta1*(tx2-MMT)+beta2*(MMT-tx2)/(tx2*MMT-
(tx2+MMT)*delta+delta^2)),'beta2')
fdel<-D(expression(beta1*(tx2-MMT)+beta2*(MMT-tx2)/(tx2*MMT-
(tx2+MMT)*delta+delta^2)),'delta')

zeta2<-beta1*(tx2-MMT)+beta2*(MMT-tx2)/(tx2*MMT-
(tx2+MMT)*delta+delta^2)
mat3<-rbind(eval(fb1),eval(fb2),eval(fdel))
mat4<-as.matrix(mat3)
se2<-sqrt(t(mat4)%*%Qvcov%*%mat4) ###se of RR

#RR diff
tx<-quantile(temp,c(0.90),na.rm=TRUE)
tx2<-quantile(temp,c(0.99),na.rm=TRUE)
fb1<-D(expression(beta1*(tx2-tx)+beta2*(tx-tx2)/(tx2*tx-
(tx2+tx)*delta+delta^2)),'beta1')   ###derivative
fb2<-D(expression(beta1*(tx2-tx)+beta2*(tx-tx2)/(tx2*tx-
(tx2+tx)*delta+delta^2)),'beta2')
fdel<-D(expression(beta1*(tx2-tx)+beta2*(tx-tx2)/(tx2*tx-
(tx2+tx)*delta+delta^2)),'delta')

zeta3<-beta1*(tx2-tx)+beta2*(tx-tx2)/(tx2*tx-(tx2+tx)*delta+delta^2)
mat5<-rbind(eval(fb1),eval(fb2),eval(fdel))
mat6<-as.matrix(mat5)
se3<-sqrt(t(mat6)%*%Qvcov%*%mat6) ###se of RR

#colnames(RRresult)<-c("MMT","RR of 90%temp","LCL","UCL","RR of 99%
```

```
temp","LCL","UCL","RR-diff","LCL","UCL")

RRresult[i,1]<-MMT
RRresult[i,2]<-exp(zeta)
RRresult[i,3]<-exp(zeta-1.96*se)
RRresult[i,4]<-exp(zeta+1.96*se)
RRresult[i,5]<-exp(zeta2)
RRresult[i,6]<-exp(zeta2-1.96*se2)
RRresult[i,7]<-exp(zeta2+1.96*se2)
RRresult[i,8]<-exp(zeta3)
RRresult[i,9]<-exp(zeta3-1.96*se3)
RRresult[i,10]<-exp(zeta3+1.96*se3)
}


######### advanced model, 3 Japan cities MV14##########

MMTmat<-matrix(NA,5,2)
colnames(MMTmat)<-c("MMT","se")
rownames(MMTmat)<-c("seoul","tokyo","osaka","nagoya","taipei")

RRresult<-matrix(NA,3,10)
rownames(RRresult)<-c("Tokyo","Osaka","Nagoya")
colnames(RRresult)<-c("RR of 90%temp","LCL","UCL","RR of 99%
temp","LCL","UCL","RR-diff","LCL","UCL")

tempadj<-c(9,1,0,1,0)

for(i in 1:5){

total<-read.csv("E:\\ms journal\\master\\totaldata.csv")
total$newcode[total$citycode == 1] <- 5
total$newcode[total$citycode == 11] <- 1
total$newcode[total$citycode == 1131] <- 2
total$newcode[total$citycode == 1271] <- 3
total$newcode[total$citycode == 1231] <- 4

total<-subset(total,newcode==i)
temp<-total$meantemp
nonacc_tot<-total$nonacc_tot
humid<-total$meanhumi
press<-total$meanpress
influ<-total$influenza_new
total$time2<-seq(1,1000,len=nrow(total))
time2<-total$time2
dow<-cbind(total$dowT,total$dowW,total$dowTh,total$dowF,total$dowS,total$dowSu)
total$time<-seq(1,nrow(total))
```

```
time<-total$time

temp<-EMA(temp,n=14)+tempadj[i]

fourier<-harmonic(time,nfreq=1,period=365.25)
nsHumid<-ns(humid,df=3)
nsPress<-ns(press,df=3)


### MLE process

fnQ<-function(theta){      ##### make the quasi likelihood function
sum(
nonacc_tot*(theta[1]+log(temp)*theta[2]+(1/(temp-
theta[4]))*theta[3]+(log(theta[5]+temp))*theta[6]
+fourier[,1]*theta[7]+fourier[,2]*theta[8]+dow[,1]*theta[9]+dow[,2]*theta[10]+d
ow[,3]*theta[11]+dow[,4]*theta[12]
+dow[,5]*theta[13]+dow[,6]*theta[14]+nsHumid[,1]*theta[15]+nsHumid[,2]*theta
[16]+nsHumid[,3]*theta[17]
+nsPress[,1]*theta[18]+nsPress[,2]*theta[19]+nsPress[,3]*theta[20]
+influ*theta[21]+time2*theta[22])

-exp(theta[1]+log(temp)*theta[2]+(1/(temp-
theta[4]))*theta[3]+(log(theta[5]+temp))*theta[6]
+fourier[,1]*theta[7]+fourier[,2]*theta[8]+dow[,1]*theta[9]+dow[,2]*theta[10]+d
ow[,3]*theta[11]+dow[,4]*theta[12]
+dow[,5]*theta[13]+dow[,6]*theta[14]+nsHumid[,1]*theta[15]+nsHumid[,2]*theta
[16]+nsHumid[,3]*theta[17]
+nsPress[,1]*theta[18]+nsPress[,2]*theta[19]+nsPress[,3]*theta[20]
+influ*theta[21]+time2*theta[22])
,na.rm=TRUE)}

startdelta<-max(temp,na.rm=TRUE)+2     ###set starting point
starttheta<-min(temp,na.rm=TRUE)-2

fitglm<-summary(glm(nonacc_tot~log(temp)+I(1/(temp-
startdelta))+I(log(starttheta+temp))
+fourier+dow+nsHumid+nsPress+influ+time2,family=quasipoisson(link=log)))

startbeta0<-fitglm$coefficients[1]
startbeta1<-fitglm$coefficients[2]
startbeta2<-fitglm$coefficients[3]
startbeta3<-fitglm$coefficients[4]
startbeta4<-fitglm$coefficients[5]
startbeta5<-fitglm$coefficients[6]
startbeta6<-fitglm$coefficients[7]
startbeta7<-fitglm$coefficients[8]
```

```
startbeta8<-fitglm$coefficients[9]
startbeta9<-fitglm$coefficients[10]
startbeta10<-fitglm$coefficients[11]
startbeta11<-fitglm$coefficients[12]
startbeta12<-fitglm$coefficients[13]
startbeta13<-fitglm$coefficients[14]
startbeta14<-fitglm$coefficients[15]
startbeta15<-fitglm$coefficients[16]
startbeta16<-fitglm$coefficients[17]
startbeta17<-fitglm$coefficients[18]
startbeta18<-fitglm$coefficients[19]
startbeta19<-fitglm$coefficients[20]

###derive estimated parameters and Hessian Matrix

fit<-optim(theta<-
c(startbeta0,startbeta1,startbeta2,startdelta,starttheta,startbeta3,startbeta4,startbeta5,
startbeta6,startbeta7,
startbeta8,startbeta9,startbeta10,startbeta11,
startbeta12,startbeta13,startbeta14,startbeta15,startbeta16,startbeta17,startbeta18,st
artbeta19)
,fnQ,hessian=TRUE,control=list(fnscale=-1,maxit=1000000))

### result
resultt<-cbind(fit$par[c(2:6)],sqrt(diag(                              solve(-
fit$hessian)[c(2:6),c(2:6)]       )))

### derive dispersion par
muhat<-exp(fit$par[1]+fit$par[2]*log(temp)+fit$par[3]*(1/(temp-
fit$par[4]))+fit$par[6]*log(temp+fit$par[5])+
+fourier[,1]*fit$par[7]+fourier[,2]*fit$par[8]
+dow[,1]*fit$par[9]+dow[,2]*fit$par[10]+dow[,3]*fit$par[11]+dow[,4]*fit$par[12]
+dow[,5]*fit$par[13]
+dow[,6]*fit$par[14]+nsHumid[,1]*fit$par[15]+nsHumid[,2]*fit$par[16]+nsHumid[
,3]*fit$par[17]
+nsPress[,1]*fit$par[18]+nsPress[,2]*fit$par[19]+nsPress[,3]*fit$par[20]
+influ*fit$par[21]+time2*fit$par[22])

pearsonresi<-(nonacc_tot-muhat)^2/(muhat)
disper<-sum(pearsonresi,na.rm=TRUE)/(length(temp)-length(fit$par))
Qvcov<-disper* solve(-fit$hessian)[c(2:6),c(2:6)]

### finding solution

beta1<-fit$par[2]
beta2<-fit$par[3]
delta<-fit$par[4]
```

```
theta<-fit$par[5]
beta3<-fit$par[6]

x3<-((((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*
(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/
(beta1+beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-
(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))^2/4)+(((1/3)*((3*
(-2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/(beta1+beta3))-
( (beta1*theta-2*beta1*delta-beta2-2*beta3*delta)^2/
((beta1+beta3)^2))))^3/27)))^(1/2)
)^(1/3))*-1*((cos(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-
(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))/(2*(((((1/27)*((2
*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)^3/
((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))^2/4)-
((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)^3/
((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))^2/4)+(((1/3)*((3*
(-2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
(beta1+beta3))-( (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3))-(sqrt(3)*sin(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-
(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))/(2*(((((1/27)*((2
*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)^3/
((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))^2/4)-
((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)^3/
((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-2*beta3*delta)* (-
2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+beta3))))^2/4)+(((1/3)*((3*
(-2*delta*theta*beta1+beta1*delta^2-beta2*theta+beta3*delta^2)/
(beta1+beta3))-( (beta1*theta-2*beta1*delta-beta2-
```

```
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3)))+( (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)/(3*(beta1+beta3)))*(-1)


MMTmat[i,1]<-x3+tempadj[i]     ##### MMT


### delta method to calculate variance of MMT

fb1<-D(expression( ((((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-( (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)^(1/3))*-1*((cos(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-( (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3))-(sqrt(3)*sin(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)* (-2*delta*theta*beta1+beta1*delta^2-
```

```
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(    (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3)))+(                                  (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)/(3*(beta1+beta3)))*(-1)),'beta1')

fb2<-D(expression(              ((((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(    (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)^(1/3))*-1*((cos(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(    (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3))-(sqrt(3)*sin(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                               (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(    (beta1*theta-2*beta1*delta-beta2-
```

```
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3)))+(                                    (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)/(3*(beta1+beta3)))*(-1)),'beta2')

fb3<-D(expression(              ((((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(   (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)^(1/3))*-1*((cos(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(   (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3))-(sqrt(3)*sin(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                                (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(   (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3)))+(                                    (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)/(3*(beta1+beta3)))*(-1)),'beta3')
```

```
fb4<-D(expression(            ((((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(  (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)^(1/3))*-1*((cos(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(  (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3))-(sqrt(3)*sin(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                    (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(  (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3)))+(                         (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)/(3*(beta1+beta3)))*(-1)),'delta')

fb5<-D(expression(            ((((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
```

```
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(   (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)^(1/3))*-1*((cos(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(   (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3))-(sqrt(3)*sin(acos(-((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))/(2*(((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)-((((1/27)*((2*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^3/((beta1+beta3)^3))-(9*(beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)*                              (-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/((beta1+beta3)^2))+(27*(beta1*delta^2*theta)/(beta1+
beta3))))^2/4)+(((1/3)*((3*(-2*delta*theta*beta1+beta1*delta^2-
beta2*theta+beta3*delta^2)/(beta1+beta3))-(   (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)^2/((beta1+beta3)^2))))^3/27)))^(1/2)
)))/3)))+(                                    (beta1*theta-2*beta1*delta-beta2-
2*beta3*delta)/(3*(beta1+beta3)))*(-1)),'theta')

mat<-rbind(eval(fb1),eval(fb2),eval(fb4),eval(fb5),eval(fb3))
mat2<-as.matrix(mat)
se3<-sqrt(t(mat2)%*%Qvcov%*%mat2) ###se of MMT
MMTmat[i,2]<-se3    ##### s.e. of MMT
}
```

Bibliography

1. Curriero, F.C., et al., *Temperature and mortality in 11 cities of the eastern United States.* American journal of epidemiology, 2002. **155**(1): p. 80-87.
2. Gasparrini, A., et al., *Mortality risk attributable to high and low ambient temperature: a multicountry observational study.* The Lancet, 2015.
3. Guo, Y., et al., *Global variation in the effects of ambient temperature on mortality: a systematic evaluation.* Epidemiology, 2014. **25**(6): p. 781-9.
4. Tobías, A., et al., *Effects of high summer temperatures on mortality in 50 Spanish cities.* Environmental Health, 2014. **13**(1): p. 48.
5. Basu, R., *Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence.* Epidemiologic Reviews, 2002. **24**(2): p. 190-202.
6. Anderson, B.G. and M.L. Bell, *Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States.* Epidemiology (Cambridge, Mass.), 2009. **20**(2): p. 205.
7. Kinney, P.L., et al., *Approaches for estimating effects of climate change on heat-related deaths: challenges and opportunities.* environmental science & policy, 2008. **11**(1): p. 87-96.
8. Patz, J.A., et al., *The potential health impacts of climate variability and change for the United States: executive summary of the report of the health sector of the US National Assessment.* Environmental health perspectives, 2000. **108**(4): p. 367.
9. Muggeo, V.M., *Estimating regression models with unknown break-points.* Stat Med, 2003. **22**(19): p. 3055-71.
10. Kim, H., J.S. Ha, and J. Park, *High temperature, heat index, and mortality in 6 major cities in South Korea.* Arch Environ Occup Health, 2006. **61**(6): p. 265-70.
11. Hastie, T. and R. Tibshirani, *Generalized additive models.* Statistical science, 1986: p. 297-310.
12. Chung, Y., et al., *Mortality related to extreme temperature for 15 cities in northeast Asia.* Epidemiology, 2015. **26**(2): p. 255-62.
13. Efron, B. and R.J. Tibshirani, *An introduction to the bootstrap*. 1994: CRC press.
14. Kunsch, H.R., *The jackknife and the bootstrap for general stationary observations.* The Annals of Statistics, 1989: p. 1217-1241.
15. Gossl, C. and H. Kuchenhoff, *Bayesian analysis of logistic regression with an unknown change point and covariate measurement error.* Stat Med, 2001. **20**(20): p. 3109-21.
16. Hoff, P.D., *A first course in Bayesian statistical methods*. 2009: Springer Science

& Business Media.
17. Box, G.E. and G.C. Tiao, *Bayesian inference in statistical analysis*. Vol. 40. 2011: John Wiley & Sons.
18. McCullagh, P., J.A. Nelder, and P. McCullagh, *Generalized linear models*. Vol. 2. 1989: Chapman and Hall London.
19. Nelder, J.A. and R. Mead, *A simplex method for function minimization.* The computer journal, 1965. **7**(4): p. 308-313.
20. Oehlert, G.W., *A note on the delta method.* The American Statistician, 1992. **46**(1): p. 27-29.
21. Papanicolaou, A., *Taylor approximation and the delta method*. 2009.
22. Casella, G. and R.L. Berger, *Statistical inference*. Vol. 2. 2002: Duxbury Pacific Grove, CA.
23. Jacobson, N., *Basic algebra I*. 2012: Courier Corporation.
24. Agresti, A., *Categorical data analysis*. 2013: John Wiley & Sons.
25. Sistrom, C.L. and C.W. Garvan, *Proportions, Odds, and Risk 1.* Radiology, 2004. **230**(1): p. 12-19.
26. Bhaskaran, K., et al., *Time series regression studies in environmental epidemiology.* Int J Epidemiol, 2013. **42**(4): p. 1187-95.
27. Gasparrini, A., B. Armstrong, and M.G. Kenward, *Distributed lag non-linear models.* Stat Med, 2010. **29**(21): p. 2224-34.
28. Armstrong, B., *Models for the relationship between ambient temperature and daily mortality.* Epidemiology, 2006. **17**(6): p. 624-31.
29. Peng, R.D., F. Dominici, and T.A. Louis, *Model choice in time series studies of air pollution and mortality.* Journal of the Royal Statistical Society: Series A (Statistics in Society), 2006. **169**(2): p. 179-203.
30. Gasparrini, A. and B. Armstrong, *Time series analysis on the health effects of temperature: advancements and limitations.* Environ Res, 2010. **110**(6): p. 633-8.
31. Hajat, S., et al., *Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London.* Epidemiology, 2005. **16**(5): p. 613-620.
32. Pattenden, S., B. Nikiforov, and B. Armstrong, *Mortality and temperature in Sofia and London.* Journal of Epidemiology and Community Health, 2003. **57**(8): p. 628-633.
33. Keatinge, W. and G. Donaldson, *Mortality related to cold and air pollution in London after allowance for effects of associated weather patterns.* Environmental Research, 2001. **86**(3): p. 209-216.
34. Härdle, W., J. Horowitz, and J.P. Kreiss, *Bootstrap methods for time series.* International Statistical Review, 2003. **71**(2): p. 435-459.
35. Politis, D.N., *The impact of bootstrap methods on time series analysis.* Statistical Science, 2003. **18**(2): p. 219-230.
36. Hastie, T.J. and R.J. Tibshirani, *Generalized additive models*. Vol. 43. 1990: CRC Press.
37. Almon, S., *The distributed lag between capital appropriations and expenditures.* Econometrica: Journal of the Econometric Society, 1965: p. 178-196.
38. Hilbe, J., *Negative binomial regression*. 2011: Cambridge University Press.

초록

서울대학교 보건대학원

보건학과 보건통계학 전공

이환희

<연구배경> 최소사망온도(Minimum mortality temperature, MMT)는 기온이 사망에 미치는 영향을 연구하는데 중요한 역학적 개념이다. 이러한 많은 연구에서 MMT를 추정하기 위해 piecewise regression나 generalized additive models(GAMs)이 많이 사용되었으나, 이 방법들은 뚜렷한 단점을 가진다. Piecewise regression은 온도와 사망 간의 비선형 관계를 반영하는데 한계를 가지고 있으며, GAMs은 MMT의 분산을 추정하는데 어려움을 가진다. 그리하여 본 연구에서는 위 두 방법의 단점을 보완하는 방법을 제시하고자 한다.

<연구방법> 본 연구에서는 MMT를 추정하기 위하여 비선형 모수를 포함한 generalized linear model을 사용하였으며, Delta Method를 사용하여 MMT의 분산을 모수적으로 추정하였다. 또한 역학에서 효과 크기 추정을 위해 사용되는 Relative Risk(RR)를 추정하였으며, 환경 연구에서 사용되는 RR간의 차이와 그 신뢰구간을 추정하였다. 본 연구에 사용된 자료는 5개 도시의 아시아 자료가 사용되었으며, 그 기간은 1992-2010 (서울), 1972-2010 (도쿄, 오사카, 나고야), and 1994-2007 (타이페이)이다.

<연구결과> 본 연구에서 제시한 새로운 모형은 온도와 사망간의 관계를 적절하게 반영하였으며, 역학적으로 합리적인 MMT를 추정하였다. 모든 도시의 MMT는 22-29° C 사에 존재하였으며, 서울을 제외하고는 0.4 이하의 표준오차를 보였으며, 연구에서 사용하는 B-spline 모형과도 유사한 결과를 보였으며, 더욱 안정적인 추정을 보였다. 나아가 폭염 연구에서 사용되는 90%, 99% 온도에서의 RR과 그 차이를 추정한 결과 또한, 기존 방법과 유사하였다.

<결론> 본 연구에서 제시된 방법은 piecewise regression 과 GAM의 단점을 보완할 수 있음을 확인하였다.